On The Robustness Of Rainfall Mapping From Measurements Of The Received Signal Level In Communication Microwave Networks

Gat S.[1], Messer H.[1]

[1]School of Electrical Engineering, Tel Aviv University, Israel.

*corresponding author: Gat Shani

e-mail: shanigat@post.tau.ac.il

Abstract

The use of existing measurements of the received signal level (RSL) at existing communication microwave networks (CMLs) for rainfall 2-D mapping has been recently shown to be a powerful tool. While previous contributions have focused on the rain mapping algorithms and their performance, in this paper we discuss their robustness. As the use of the RSL of microwave links as virtual sensors for rainfall monitoring is an unintended one, the signals are not optimized for best sensing conditions, and the signal processing algorithms need to deal with practical problems. Here we focus on the problem of outliers in the RSL measurements, which can be caused by technical failure or other reasons. We analyze the effect of outliers on the standard rain mapping algorithms, showing that the improved algorithm proposed by Goldstein *et al*. has inherent robustness. Then, we propose to modify it to further improve its robustness and we demonstrate how. The results are demonstrated with real data and simulations.

Keywords: Spatial interpolation, rainfall mapping, robustness, microwave links

1. Introduction

The idea of using existing measurements of signal levels from commercial microwave links (CMLs) for environmental monitoring was first introduced in Messer *et al*, 2006, 2007 and has been spread since Goldstein and Messer, 2015. In particular, the availability of many, spatially distributed, virtual rainfall sensors enables near-ground rainfall mapping. Different rainfall 2-D mapping algorithms, based on existing CMLs data, have been proposed (Goldstein *et al*, 2009, and Overeem *et al*, 2015, 2016). Many of them are based on translating links' attenuation to an average rain rate along the link, using the well-known power law (Olsen *et al*, 1978):

$$A = aR^b \qquad (1)$$

Where A is the attenuation (dB per km) along the transmission path, R is the rain-rate (mm/hr), and a and b are frequency dependent constants. The average rain rate is used as if it was measured by a rain gauge (or gauges) in the link's general location (usually, its middle, or as separate "gauges" along its length). That is, each link is converted to virtual rain gauges using (1). Those "virtual gauges" measurements are interpolated using a common interpolation scheme, such as inverse distance weighting (IDW), Kriging, and similar methods (Goldstein *et al*, 2009, and Overeem *et al*, 2016). These algorithms have not dealt directly with robustness to outliers, which may be a specific issue when using CMLs due to technological failures. In general, a common way to avoid inaccuracies due to data outliers is to apply a pre-processing method excluding extreme values (Overeem *et al*, 2016)., such as excluding a certain percentage of the measurements in a predefined time span. In this paper we consider generic spatial rain mapping from a single snapshot. That is, neither prior knowledge on the tempo-spatial behavior of rainfall is assumed, nor training is required. The paper is organized as follows: in Section 2 we demonstrate the need for a robust algorithm by showing potential outliers in real data. In Section 3, a modification to existing algorithms which improves robustness will be developed. Section 4 will feature some simulated results for the robust algorithm, followed by some concluding remarks.

2. Experimental Evidence For Outliers

As an example, we use measurements collected in 2015 in Sweden within a pilot collaboration between Ericsson and SMHI 0). The RSL data was translated to an average rain-rate along each link using (1), as described in 0) after subtracting the attenuation base-line (Zinevich *et al*, 2010) using standard method. In Figure 1, 160 CMLs in Gothenburg's area can be seen. Each quadrant is about 10 km^2, an area over which rain doesn't change much. As such, we expect the average rain-rate sensed by all links is each quadrant to be similar. A histogram of the average rain rate along them for two quadrants of the map is given in Figure 2,3 The data presented taken for 15:32 UTC on June 2ed, 2015. The distribution of rain-rate over the south-eastern quadrant is rather smooth, between 3-17 mm/hr of rain-rate, with two links indicating stronger rain that may or may not be outliers. However, an extreme value of 60 mm/hr in one of the links at the north-eastern quadrant definitely indicates an outlier (this outlier CML located at the red dot in Figure 1, in close proximity to others). The influence of extreme measurements (outliers) on the rainfall mapping depends on the spatial interpolation method chosen. In the following we will
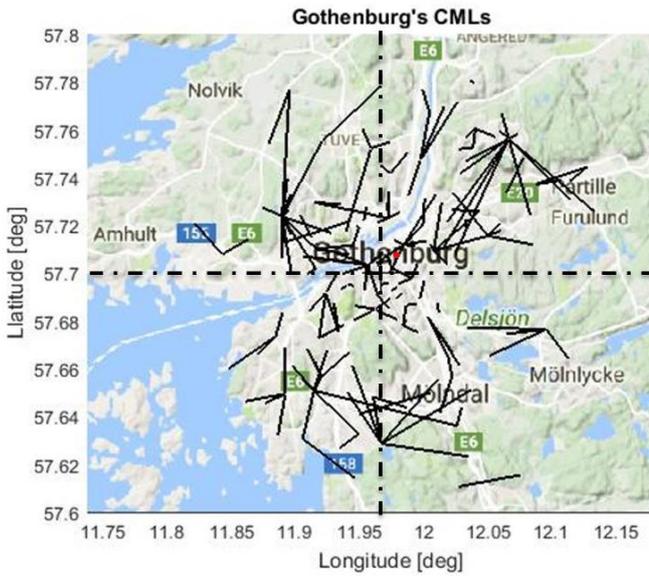
Figure 1: A map of Gothenburg's area and CMLs' locations.
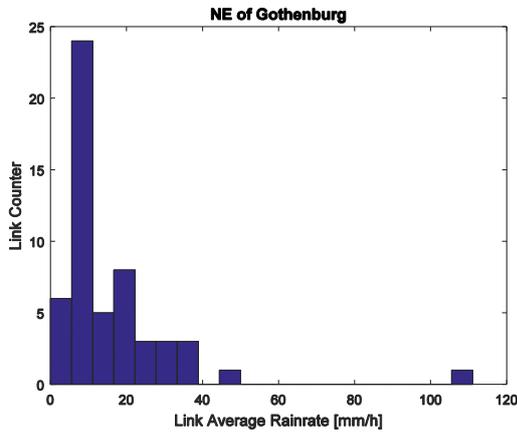
Figure 2: Histogram of average rain rates along the links shown in Figure 1. Links of the North-Eastern quadrant (total of 54 CMLs)
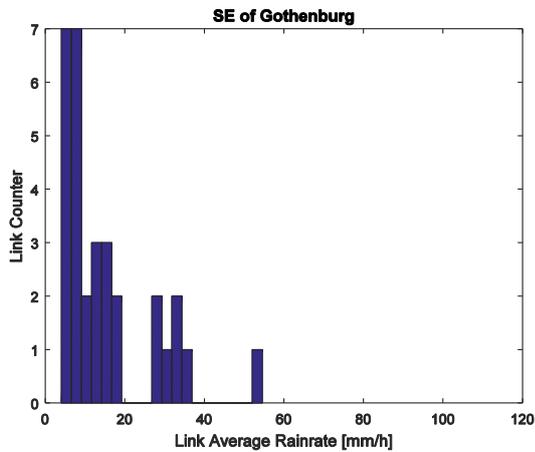
Figure 3: Histogram of average rain rates along the links shown in Figure 1. Links of the South-Eastern quadrant (total of 31 CMLs).

focus on the method described in Goldstein *et al*, 2009.

3.   The Robust Approach

In Goldstein *et al*, 2009, the attenuation of the links, (following a pre-processing to retrieve rain-related attenuation only) is translated to a series of 'virtual gauges' along the link. The initial assigned value to these gauges is the "naïve" average rain-rate given by the power-law (1). Then, the IDW interpolation method has been used for interpolating between the virtual rain gauges. There, it has been shown that this method is equivalent to a weighted least square (WLS) with an appropriate noise covariance matrix. While for the classic IDW the noise covariance matrix is assumed diagonal, with CML's noise variance which is inversely proportional to the distance between the virtual gauge and the estimated point, Goldstein *et al*. also suggested an improved WLS where the noise covariance matrix also consider the quantization noise at each link. Moreover, the algorithm in Goldstein *et al*, 2009 does not assume that the rain is constant along a link and it iteratively estimates the rain rate measured by virtual rain gauges along the microwave link, based on the interpolation results. The iterative estimation consists of two steps, for each link:

Step 1: Interpolated estimation of rain rate values at the virtual gauges using data measurements from all other links.

Step 2: Correcting step 1's results using link's actual measurements as a constraint.

Keeping with the paper's notations, the iteration's results for the ith link are given by minimizing (2) below with λ being a Lagrange multiplier:

$$F_i = \sum_{j=1}^{K_i} \left(r_{ij}^b - \hat{\theta}_{ij}^b \right)^2 + \lambda_i \left(\sum_{j=1}^{K_i} r_{ij}^b - K_i R_i^b \right) \quad (2)$$

Where r_{ij} describes the virtual gauges measurements, $\hat{\theta}_{ij}$ the step 1 estimation result, K_i the number of virtual gauges along the link and b the frequency dependent power of (1), and given in ITU, 2005. Minimizing (2) leads to:

$$r_{ij} = \left(R_i^b - K_i^{-1} \sum_{j=1}^{K_i} \hat{\theta}_{ij}^b + \hat{\theta}_{ij}^b \right)^{\frac{1}{b}} \quad (3)$$

The $R_i^b - K_i^{-1} \sum_{j=1}^{K_i} \hat{\theta}_{ij}^b$ term is the correction derived directly from the link's measured attenuation, being a constraint. Clearly, the constraint's argument has great effect on the estimation of the rain rate in a virtual gauge - its effect equals the combined effect of all the other links' measurements, represented by $\hat{\theta}_{ij}$. Therefore, adding a large negative constraint may result in a negative value within the brackets (and an impossible complex rain rate value for the virtual gauge). The algorithm's solution in0 is simply assuming the value to be zero for such cases. But the histogram in Figure 2 suggests that the use of this constraint is not robust, as it will force an erroneous value for the outlier link's virtual gauges. Therefore a more complete approach should be taken. Hence, abandoning the

constraint feature and replacing it with a more general function is suggested:

$$\tilde{F}_i = \sum_{j=1}^{K_i}\left(r_{ij}^b - \hat{\theta}_{ij}^b\right)^2 + \rho\left(\sum_{j=1}^{K_i} r_{ij}^b - K_i R_i^b\right) \quad (4)$$

The new function $\rho(x)$, (in this case $x = \sum_{j=1}^{K_i} r_{ij}^b - K_i R_i^b$) should balance the influence of the measurements of neighboring links, and the measurements of the estimated link itself. Note that for $\rho(x) = \lambda_i x$, (4) equals (2). A robust $\rho(x)$ is usually a symmetric function that behaves as quadratic function near zero, and changes to a lass steep function for larger $|x|$ (Bai, 2012, Huber, 1981). Differentiation of (4) with respect to r_{ij}^b yields:

$$2\left(r_{ij}^b - \hat{\theta}_{ij}^b\right) + \rho'\left(\sum_{j=1}^{K_i} r_{ij}^b - K_i R_i^b\right) = 0 \quad (5)$$

We notice that for $x > \hat{\theta}_{ij}^b$, the expression in the parentheses in (3) is negative. In order to get solution (3) for function $\rho(x)$ in (5), the Goldstein- algorithm actually applies:

$$\rho'(x) = \begin{cases} 2\theta, & x > \theta \\ \lambda_i, & x \le \theta \end{cases}, \text{ therefore:}$$

$$\rho(x) = \begin{cases} 2\theta x, & x > \theta \\ \lambda_i x, & x \le \theta \end{cases} \quad (6)$$

Initially, x is unknown (because r_i is unknown). A guess to evaluate x will be to use some x_0. Hence:

$$\rho(x_0, x) = \begin{cases} 2\theta x, & x_0 > \theta \\ \lambda_i x, & x_0 \le \theta \end{cases} \quad (7)$$

The x_0 used is $x_0 = K_i^{-1}\sum_{j=1}^{K_i} \hat{\theta}_{ij}^b - R_i^b$.

This is a non-symmetric constant-quadratic function. An obvious initial 'fix' to a symmetric constraint will be switching to a symmetric constant-quadratic-constant function:

$$\tilde{\rho}(x_0, x) = \begin{cases} 2\theta x & x_0 \le -\theta \\ \lambda_i x, & elsewhere \\ 2\theta x & x_0 \ge \theta \end{cases} \quad (8)$$

Or in other words, the estimation is winsorized (Huber, 1981) to the span: $r_{ij}^b \in [0, 2\hat{\theta}_{ij}^b]$.

4. Simulations

A simulation using the actual links' distribution of Gothenburg's area was used in order to verify the increased robustness of the suggested ρ function (7). The simulated rainfall profile of Figure 1 was translated to attenuation for each link using (1), numerically integrating the attenuation along all links, with 10dB addition to 10% of the links, and chosen randomly, as possible extreme outliers. Since the RSL data retrieved from real CMLs is

quantized, the simulated RSL was quantized as well, using $\Delta A = 0.3dB$ as the quantization resolution. The simulated links' data was used to estimate the rain rate at the virtual gauges was estimated and interpolated, using the algorithm of Goldstein *et al*, 2009, and the symmetric (and potentially more robust) version described in part 3. Some results are shown in Figures 6,7 and in Table 1. The runs differ only in the links chosen for the addition of outlier noise. The "true" rain profile and the links' distribution and properties remain the same for all runs.
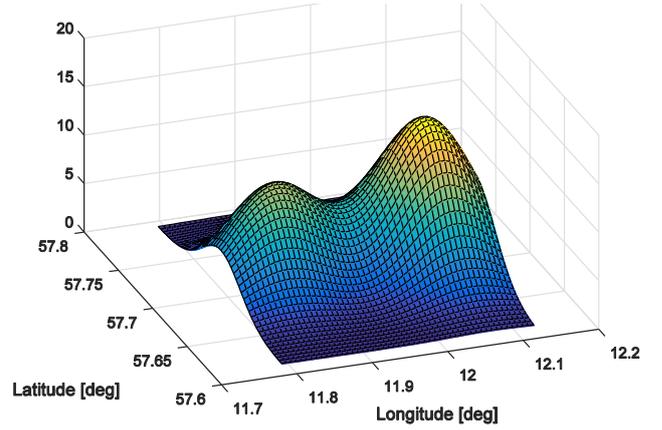
Figure 4: Map of simulated rainfall. A 3-D map, height and colors indicate rain-rate [mm/h].
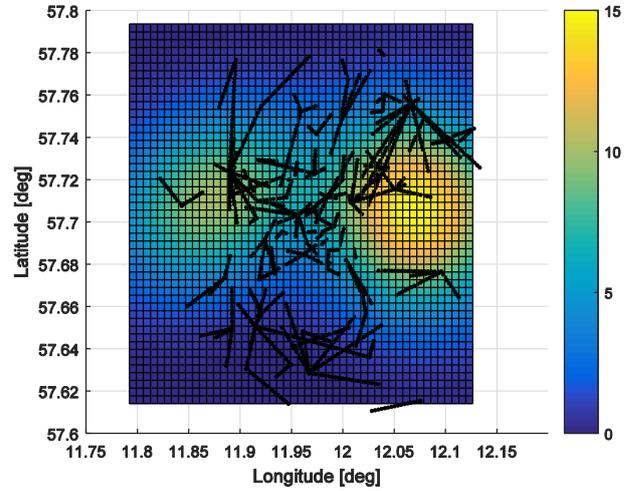
Figure 5: Map of simulated rainfall. A 2-D map, colors indicate rain-rate [mm/h].

Table 1 shows the correlation and MSE (mean-square-error) of the true rain-rate map versus the estimated rain-rate map. The symmetric algorithm shows consistently similar or better correlation to the ground truth, and smaller MSE. The upper map in Figure 7 is an example of the effect of possible outliers: the south-west corner of the estimation map shows strong rain where there was none.

5. Discussion And Conclusion

Table 1: Simulation results

run	The Original Algorithm (Goldstein *et al*. 2009)		Modified Robust Algorithm	
	Corr.	**MSE**	**Corr.**	**MSE**
1	0.60	21.09	0.82	5.88
2	0.51	22.91	0.82	6.20
3	0.83	14.40	0.87	5.08
4	0.81	9.33	0.86	4.48
5	0.74	19.39	0.77	9.76
Mean	**0.70**	**17.42**	**0.83**	**6.28**
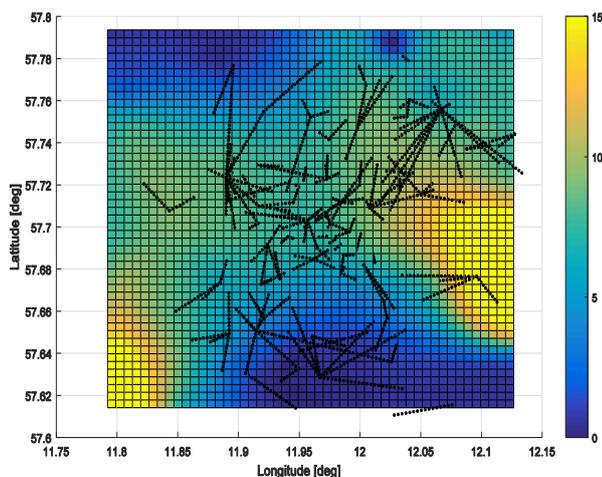
Figure 6: 2-D maps of estimated rain-rate. Results for the original Goldstein-algorithm.
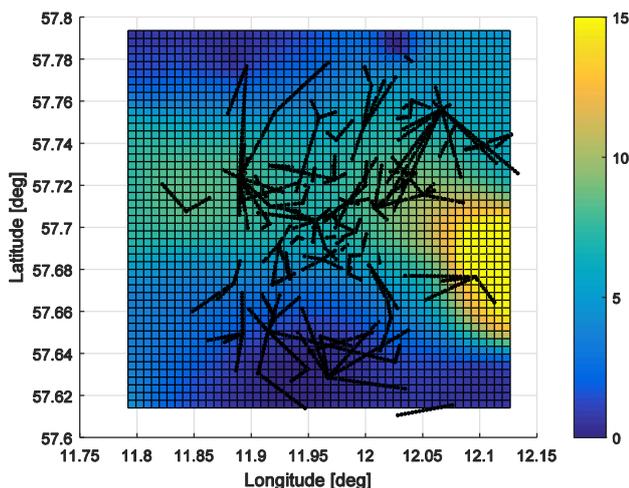
Figure 7: 2-D maps of estimated rain-rate. Results for suggested symmetric robust algorithm.

In this paper we have presented real data and simulation results indicating the need to improve robustness in 2-D rainfall mapping algorithms from CMLs. The real data, taken from an urban area with high spatial (160 links for a 20 km^2 area) and temporal (measurement every 10 seconds) resolution, allows the detection of outliers in measurements with good probability. Simulations of randomly chosen defective measurements for a single scenario showed the potential vulnerability of existing methods. We suggest an initial 'fix', based on common methods of robust regression. But, since this is an ad-hoc solution, its optimality is not guaranteed. A more optimal algorithm should be possible to find, based on the rainfall probability distribution. A possible approach may be to leave the ad-hoc definition of the minimization of (2), and instead, assume a statistical relation between measurements in different locations (the virtual gauges) and the RSL measurements, allowing the use of the maximum-likelihood approach, or robust variants of M-estimation (Gat, 2017).

References

Bai, X., "Robust linear regression "(Doctoral dissertation,

Kansas State University) , 2012

Gat, Shani. *Robust Estimation and Mapping of Rainfall from*

Microwave Measurements. MSc thesis, Tel-Aviv University, 2017.

Goldshtein, Oren, Hagit Messer, and Artem Zinevich. "Rain rate

estimation using measurements from commercial telecommunications links." *IEEE Transactions on Signal Processing* 57.4 (2009): 1616-1625.

Huber, P.J., "Robust statistics", 1981.

Messer, Hagit, Artem Zinevich, and Pinhas Alpert.

"Environmental monitoring by wireless communication networks." *Science* 312.5774 (2006): 713-713.

Messer, Hagit. "Rainfall Monitoring Using Cellular Networks

[In the Spotlight]." *IEEE Signal Processing Magazine* 24.3 (2007): 144-142.

Messer, Hagit, and Oren Sendik. "A New Approach to

Precipitation Monitoring: A critical survey of existing technologies and challenges." *IEEE Signal Processing magazine* 32.3 (2015): 110-122

Olsen, Roderic, David Rogers, and Daniel Hodge. "The aR b

relation in the calculation of rain attenuation." *IEEE Transactions on Antennas and Propagation* 26.2 (1978): 318-329.

Overeem, Aart, Hidde Leijnse, and Remko Uijlenhoet.

"Country-wide rainfall maps from cellular communication networks." *Proceedings of the National Academy of Sciences* 110.8 (2013): 2741-2745.

Overeem, Aart, Hidde Leijnse, and Remko Uijlenhoet.

"Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network." *Atmospheric Measurement Techniques* 9.5 (2016): 2425-2444.

Rec. ITU-R P.838-3 1, (2005)

Zinevich, Artem, Hagit Messer, and Pinhas Alpert. "Prediction

of rainfall intensity measurement errors using commercial microwave communication links." *Atmospheric Measurement Techniques* 3.5 (2010): 1385.

https://www.ericsson.com/mobility-report/microweather-unlocking-potential